

Before and below ‘theory of mind’: embodied simulation and the neural correlates of social cognition

Vittorio Gallese*

Department of Neuroscience, Section of Physiology, University of Parma, 43100 Parma, Italy

The automatic translation of folk psychology into newly formed brain modules specifically dedicated to mind-reading and other social cognitive abilities should be carefully scrutinized. Searching for the brain location of intentions, beliefs and desires—as *such*—might not be the best epistemic strategy to disclose what social cognition really is. The results of neurocognitive research suggest that in the brain of primates, mirror neurons, and more generally the premotor system, play a major role in several aspects of social cognition, from action and intention understanding to language processing. This evidence is presented and discussed within the theoretical frame of an embodied simulation account of social cognition. Embodied simulation and the mirror neuron system underpinning it provide the means to share communicative intentions, meaning and reference, thus granting the parity requirements of social communication.

Keywords: embodied simulation; folk psychology; mind-reading; mirror neurons; monkey; theory of mind

1. INTRODUCTION

The traditional view in the cognitive sciences holds that humans are able to understand the behaviour of others in terms of their mental states—intentions, beliefs and desires—by exploiting what is commonly designated as ‘folk psychology’. According to a widely shared view, non-human primates, including apes, do not rely on mentally based accounts of each other’s behaviour.

This view prefigures a sharp distinction between non-human species, confined to behaviour-reading, and our species, whose social cognition makes use of a different level of explanation, i.e. mind-reading. However, it is by no means obvious that behaviour-reading and mind-reading constitute two autonomous realms. In fact, during our social transactions, we seldom engage in explicit interpretative acts. Most of the time, our understanding of social situations is immediate, automatic and almost reflex-like. Therefore, it seems preposterous to claim that our capacity to reflect on the intentions, beliefs and desires determining the behaviour of others is *all there is* in social cognition. It is even less obvious that, while understanding the intentions of others, we employ a cognitive strategy totally unrelated to predicting the consequences of their observed behaviours. A growing sense of discomfort towards a blind faith in folk psychology to characterize social cognition is indeed surfacing within the field of philosophy of mind. It has recently been stressed that the use of folk psychology in social cognition of the belief-desire propositional attitudes is overstated (see Hutto 2004). As emphasized by Bruner (1990, p. 40), ‘when

things are as they should be, the narratives of folk psychology are unnecessary’.

Another problem for the mainstream view on social cognition is posed by the relationship between mind-reading and linguistic competence. Recent evidence shows that 15-month-old infants understand false beliefs (Onishi & Baillargeon 2005). These results suggest that typical aspects of mind-reading, like the attribution of false beliefs to others, can be explained on the basis of low-level mechanisms which develop well before full-blown linguistic competence.

The point I want to stress is that social cognition is not only ‘social metacognition’; that is, *explicitly* thinking about the contents of someone else’s mind by means of symbols or other representations in propositional format. We can certainly ‘explain’ the behaviour of others by using our complex and sophisticated mentalizing abilities. And we should add that the neural mechanism underpinning such complex mentalizing abilities are far from being fully understood. Most of the time, though, we do not need to do this. We have a much more direct access to the inner world of others. Direct understanding does not require explanation. This particular dimension of social cognition is embodied, in that it mediates between the multimodal experiential knowledge of our own lived body and the way we experience others.

I have presented elsewhere the accounts of how embodied simulation can underpin basic forms of social cognition like the capacity of empathizing with others’ emotions and sensations (Gallese 2001, 2003*a,b*, 2005*a,b*). The main goal of the present article is more ambitious. It is to show that embodied simulation can play an explanatory role not only on low-level mechanisms of social cognition—like those involved in empathy—but also on its more sophisticated

*vittorio.gallese@unipr.it

One contribution of 19 to a Discussion Meeting Issue ‘Social intelligence: from brain to culture’.

aspects—like the attribution of mental states to others, and language. For this purpose, I briefly summarize the functional properties of the mirror neuron system in monkeys and humans. I show that this system is involved in different aspects of social cognition like action and intention understanding and social communication. I also show that the premotor system is at the basis of different aspects of the faculty of language. I conclude by introducing that the ‘neural exploitation hypothesis’, according to which a single functional mechanism, embodied simulation, is probably at the basis of various and important aspects of social cognition.

2. THE MIRROR NEURON SYSTEM FOR ACTIONS IN MONKEYS AND HUMANS

More than a decade ago, a new class of motor neurons, mirror neurons, was discovered in area F5 within the ventral premotor cortex of the macaque monkey. These neurons discharge not only when the monkey executes goal-related hand and/or mouth acts like grasping objects, but also when observing other individuals (monkeys or humans) executing similar actions (di Pellegrino *et al.* 1992; Gallese *et al.* 1996; Rizzolatti *et al.* 1996; Ferrari *et al.* 2003). Neurons with similar mirroring properties, matching action observation and execution, have also been discovered in a sector of the posterior parietal cortex reciprocally connected with area F5 (see Rizzolatti *et al.* 2001; Gallese *et al.* 2002; Fogassi *et al.* 2005). It has been proposed that this ‘direct matching’ may underpin a direct form of action understanding (Gallese *et al.* 1996; Rizzolatti *et al.* 1996, 2001; Gallese *et al.* 2004; Rizzolatti & Craighero 2004) by exploiting embodied simulation, a specific mechanism by means of which the brain/body system models its interactions with the world (Gallese 2001, 2003a,b, 2005a,b, 2006).

In order to test the hypothesis that mirror neurons underpin action understanding via embodied simulation, we assessed their activation in conditions in which the monkey understands the meaning of the occurring action, but has no access to the visual features that activate mirror neurons. If mirror neurons really underpin action understanding, their activity should reflect the meaning of the observed action rather than its visual features. Experiments by Umiltà *et al.* (2001) showed that F5 mirror neurons become active also during the observation of partially hidden actions, when the monkey can predict the action outcome, even in the absence of the complete visual information about it (Umiltà *et al.* 2001). Macaque monkey’s mirror neurons therefore map actions made by others not just on the basis of their visual description, but also on the basis of the anticipation of the final goal of the action, by means of the activation of its motor representation in the observer’s premotor cortex.

In another series of experiments, we showed that a particular class of F5 mirror neurons (‘audio–visual mirror neurons’) respond not only when the monkey executes and observes a given hand action, but also when it just hears the sound typically produced by the action (Kohler *et al.* 2002). These neurons respond to the sound of actions and discriminate between the

sounds of different actions, but do not respond to other similarly interesting sounds. In sum, the different modes of presentation of events intrinsically different, as sounds, images or willed motor acts, are nevertheless bound together within a simpler level of semantic reference, underpinned by the same network of audio–visual mirror neurons. The presence of such a neural mechanism within a non-linguistic species can be interpreted as the neural correlate of the dawning of a conceptualization mechanism (Gallese 2003c; Gallese & Lakoff 2005).

Different experimental methodologies and techniques have also demonstrated in the human brain the existence of a mirror neuron system matching action perception and execution. During action observation, there is a strong activation of premotor and parietal areas, the probable human homologue of the monkey areas in which mirror neurons were originally described (for review, see Rizzolatti *et al.* 2001; Gallese 2003a,b, 2006; Gallese *et al.* 2004; Rizzolatti & Craighero 2004). The mirror neuron system in humans is somatotopically organized, with distinct cortical regions within the premotor and posterior parietal cortices being activated by the observation/execution of mouth-, hand- and foot-related actions (Buccino *et al.* 2001). More recently, it has been shown that the mirror neuron system in humans is directly involved in the imitation of simple finger movements (Iacoboni *et al.* 1999), as well as in learning previously never-practised complex motor acts (Buccino *et al.* 2004b).

A recent study by Buxbaum *et al.* (2005) on posterior parietal neurological patients with ‘ideomotor apraxia’ has shown that they were not only disproportionately impaired in the imitation of transitive gestures, when compared with intransitive gestures, but also showed a strong correlation between imitation deficits and the incapacity of recognizing observed goal-related meaningful hand actions. These results further corroborate the notion that the same action representations underpin both action production and action understanding.

3. THE MIRROR NEURON SYSTEM FOR COMMUNICATIVE ACTIONS IN MONKEYS AND HUMANS

The macaque monkey premotor area F5 also contains neurons related to mouth actions. In the most lateral part of area F5, we described a population of mirror neurons mostly related to the execution/observation of mouth-related actions (Ferrari *et al.* 2003). The majority of these neurons discharge when the monkey executes and observes transitive object-related ingestive actions, such as grasping, biting or licking. However, a small percentage of mouth-related mirror neurons discharge during the observation of communicative facial actions performed by the experimenter in front of the monkey (‘communicative mirror neurons’; Ferrari *et al.* 2003). These actions are affiliative gestures like lip-smacking and lips or tongue protrusion. A behavioural study showed that the observing monkeys correctly decoded these and other communicative gestures performed by the experimenter in front of them, because they elicited congruent expressive reactions (Ferrari *et al.* 2003). Communicative

mirror neurons could be an evolutionary precursor of social communication mediated by facial gestures.

A recent brain-imaging study, in which human participants observed mouth actions performed by humans, monkeys and dogs (Buccino *et al.* 2004a), corroborates this hypothesis. The observed mouth actions could be either object-directed, like a human, monkey or dog biting a piece of food, or communicative, like human silent speech, monkey lip-smacking and dog barking. The results showed that the observation of all biting actions led to the activation of the mirror neuron system, encompassing the posterior parietal and ventral premotor cortices (Buccino *et al.* 2004a). Interestingly, the observation of communicative mouth actions led to the activation of different cortical foci according to the different observed species. The observation of human silent speech activated the pars opercularis of the left inferior frontal gyrus, the premotor sector of Broca's region. The observation of monkey lip-smacking activated a smaller part of the same region bilaterally. Finally, the observation of the barking dog activated only extra-striate visual areas.

Actions belonging to the motor repertoire of the observer (e.g. biting and speech-reading) or very closely related to it (e.g. monkey's lip-smacking) are mapped on the observer's motor system. Actions that do not belong to this repertoire (e.g. barking) are mapped and, henceforth, categorized on the basis of their visual properties. These results show two things. First, the activation of the mirror neuron system is proportionate to the degree of congruence between the observed actions and the observer's motor repertoire (see also Calvo-Merino *et al.* 2005). Second, embodied simulation is not the only mechanism mediating action understanding. What I take to be crucially different between the understanding mediated by embodied simulation and that mediated by the cognitive interpretation of a visual scene (as in the case of the observed barking dog) is the quality of the experience coupled with the understanding. Only the embodied simulation mediated by the activation of the mirror neuron system enables the capacity of knowing 'how it feels' to perform a given action. Only this mechanism enables intentional attunement with the observed agent (Gallese 2006).

The involvement of the motor system during observation of communicative mouth actions is also testified by the results of a transcranial magnetic stimulation (TMS) study by Watkins *et al.* (2003), in which they showed that the observation of silent speech-related lip movements enhanced the size of the motor-evoked potential in lip muscles. This effect was lateralized to the left hemisphere. Consistent with the brain-imaging data of Buccino *et al.* (2004a), the results of Watkins *et al.* (2003) show that the observation of communicative, speech-related mouth actions facilitates the excitability of the motor system involved in the production of the same actions.

4. THE MIRROR NEURON SYSTEM FOR ACTIONS AND THE UNDERSTANDING OF INTENTIONS

What does the presence of mirror neurons in different species of primates such as macaques and humans tell us about the evolution of social cognition? The

evidence collected so far seems to suggest that the mirror neuron system for actions is sophisticated enough to enable its exploitation for social purposes. This matching mechanism indeed supports social facilitation in monkeys. It has recently been shown that the observation and hearing of noisy eating actions facilitates eating behaviour in pigtailed macaque monkeys (Ferrari *et al.* 2005).

Another recently published study shows that the pigtailed macaque monkeys recognize when they are imitated by a human experimenter (Paukner *et al.* 2005). The pigtailed macaques preferentially look at an experimenter imitating the monkeys' object-directed actions when compared with an experimenter manipulating an identical object, but not imitating their actions. Since both experimenters acted in synchrony with the monkeys, the monkeys based their gaze preference not on temporal contingency, but evidently took into account the structural components of the experimenters' actions.

Even if it is true, as repeatedly stated, that macaque monkeys are not capable of *motor* imitation—though recent evidence by Subiaul *et al.* (2004) shows that they are capable of *cognitive* imitation—the study by Paukner *et al.* (2005) nevertheless shows that macaque monkeys do entertain the capacity to discriminate between very similar goal-related actions on the basis of their degree of similarity with the goal-related actions the monkeys themselves have just executed. This capacity appears to be cognitively sophisticated, because it implies a certain degree of metacognition in the domain of purposeful actions.

But monkeys do not entertain the full-blown mentalization typical of humans. Thus, since both species do have mirror neurons, what makes humans different? The easiest answer is, of course, the presence of language. This answer, though, is at least partly question-begging, because it only transposes the human cognitive endowment to be explained. Furthermore, it implies a perfect overlap between language and our mentalizing abilities. A discussion of this debated issue is beyond the scope and space limits of this article, but I will come back to the issues of language and the evolution of social cognition in the final sections.

At present, we can only make hypotheses about the relevant and still poorly understood neural mechanisms underpinning the mentalizing abilities of humans. In particular, we do not have a clear neuroscientific model of how humans understand the intentions promoting the actions of others they observe. When an individual starts a movement aimed to attain a goal, such as picking up a pen, they have clearly in mind what they are going to do, for example writing a note on a piece of paper. In this simple sequence of motor acts, the final goal of the whole action is present in the agents' mind and is somehow reflected in each motor act of the sequence. The action *intention*, therefore, is set before the beginning of the movements. This also means that when we are going to execute a given action, we can also predict its consequences.

However, in social contexts, a given act can be originated by very different intentions. Suppose, one sees someone else grasping a cup. Mirror neurons for grasping will most probably be activated in the

observer's brain. A simple motor equivalence between the observed act and its motor representation in the observer's brain, though, can only tell us *what* the act is (it is a grasp) and not *why* it occurred. This has led us to argue against the relevance of mirror neurons for social cognition and, in particular, for determining the intentions of others (see Jacob & Jeannerod 2005).

We should ask ourselves the following question: what does it mean to determine the intention of the action of someone else? I propose a deflationary answer. Determining *why* a given act (e.g. grasping a cup) was executed can be equivalent to detecting the goal of the still not executed and impending subsequent act (e.g. bringing the cup to the mouth).

These issues were experimentally addressed with a functional magnetic resonance imaging (fMRI) study (Iacoboni *et al.* 2005). Volunteers watched three kinds of stimuli: hand grasping acts without a context; context only (a scene containing objects); and hand grasping acts embedded in contexts. In the latter condition, the context suggested the intention associated with the grasping (either drinking or cleaning up). The observation of motor acts embedded in contexts, compared with the other two conditions, yielded a significant signal increase in the posterior part of the inferior frontal gyrus and the adjacent sector of the ventral premotor cortex, where hand actions are represented. Thus, premotor mirror areas—areas active during the execution and the observation of action—previously thought to be involved only in action recognition—are actually also involved in understanding the 'why' of action, i.e. the intention of promoting it. These results suggest that for simple actions such as those employed in this study, the ascription of intentions occurs by default and it is underpinned by the mandatory activation of an embodied simulation mechanism (Gallese 2006; see also Gallese & Goldman 1998).

The neurophysiological mechanism at the basis of the relationship between intention detection and action prediction was recently clarified. Fogassi *et al.* (2005) described a class of parietal mirror neurons whose discharge during the observation of an act (e.g. grasping an object) is conditioned by the type of not-yet-observed subsequent act (e.g. bringing the object to the mouth), specifying the overall action intention. This study shows that parietal mirror neurons discharge in association with the execution/observation of motor acts (grasping) only when they are embedded in a specific action aimed at a more specific distal goal. It must be emphasized that the neurons discharge before the monkey itself executes, or observes the experimenter starting, the second motor act (bringing the object to the mouth or placing it into the cup). Single motor acts are dependent on each other, as they participate in the overarching distal goal of an action, thus forming pre-wired intentional chains, in which each subsequent motor act is facilitated by the previously executed one.

This suggests that in addition to recognizing the goal of the observed motor act, mirror neurons allow the observing monkey to predict the agent's next act, henceforth the action overall intention. This mechanism can also be interpreted as the precursor of

more sophisticated intention understanding abilities, such as those characterizing our species.

The mechanism of intention understanding just described appears to be rather simple, i.e. depending on which motor chain is activated, the observer is going to activate the motor schema of what, most probably, the agent is going to do. How can such a mechanism be formed? The statistical frequency of act sequences, as they are habitually performed or observed in the social environment, could constrain preferential paths of act inferences/predictions. This could be accomplished by chaining together different motor schemata. At the neural level, this would be equivalent to the chaining of different populations of mirror neurons coding not only the observed motor act, but also those that would normally follow in a given context.

Ascribing intentions would therefore consist in predicting a forthcoming new goal. According to this perspective, action prediction and the ascription of intentions are related phenomena, underpinned by the same functional mechanism, i.e. embodied simulation. In contrast with what mainstream cognitive science would maintain, action prediction and the ascription of intentions—at least of simple intentions—do not appear to belong to different cognitive realms, but are both related to embodied simulation mechanisms underpinned by the activation of chains of logically related mirror neurons.

The neuroscientific evidence presented so far shows that our brains, as well as those of macaques, have developed a basic functional mechanism, embodied simulation, which can provide a direct access to the meaning of the actions and intentions of others. This evidence suggests that many aspects of social cognition are tractable at the neural level of description. Let us now examine to what extent the embodied simulation account of social cognition can also be applied to the most distinctive aspect of human social cognition, i.e. language.

5. SOCIAL COGNITION AND LANGUAGE

Any account of human social cognition cannot get away from language. Language is the most specific hallmark of what it means to be human. The search for where and how language evolved and the study of the functional mechanisms at the basis of the language capacity become toolkits to explore human nature. In spite of a very long history of studies and speculations, the intimate nature of language and the evolutionary process producing it still remain somewhat elusive. One reason for such elusiveness stems from the complexity and multidimensional nature of language. What do we refer to when we investigate the language faculty and its evolution? Is language the outcome of a dedicated system, or does it include more general cognitive abilities?

What can a neuroscientific perspective add to such a controversial debate, and how can it help in clarifying social cognition? A possible starting point is to consider the fact that human language for most of its history has been just spoken language. This may suggest that language most probably evolved in order to provide individuals with a more powerful and flexible social cognitive tool to share, communicate and exchange

knowledge (Tomasello *et al.* 2005). According to this perspective, the social dimension of language becomes crucial for its understanding.

In §§6–8, I will address the issue of the relation among the faculty of language, action and embodied simulation. I will show that when processing language, humans show activation of the motor system. This activation occurs at different levels. The first level can be defined as 'motor simulation at the vehicle level', and pertains to the phono-articulatory aspects of language. The second level can be defined as 'motor simulation at the content level', and concerns the semantic content of a word, verb or proposition. Finally, I will briefly touch upon the topic of syntax.

6. EMBODIED SIMULATION AND LANGUAGE: MOTOR SIMULATION AT THE VEHICLE LEVEL

Broca's region, traditionally considered as an exclusive speech production area, contains representations of orofacial gestures and hand actions, and it is known to be part of the mirror neuron system (for review, see Bookheimer 2002; Rizzolatti & Craighero 2004; Nishitani *et al.* 2005). In a TMS experiment, Fadiga *et al.* (2002) showed that listening to phonemes induces an increase of motor-evoked potentials (MEPs) amplitude recorded from the tongue muscles involved in their execution. This result was interpreted as an acoustically related resonance mechanism at the phonological level. These results have been complemented by a TMS study of Watkins *et al.* (2003), who showed that listening to and viewing speech gestures enhanced the amplitude of MEPs recorded from the lip muscles. An activation of motor areas devoted to speech production during passive listening to phonemes has recently also been demonstrated in an fMRI study (Wilson *et al.* 2004). Finally, Watkins & Paus (2004) showed that during auditory speech perception, the increased size of the MEPs obtained by TMS over the face area of the primary motor cortex correlated with cerebral blood flow increase in Broca's area.

It is worth noting that not only speech perception, but also covert speech activates phono-articulatory simulation within the motor system. McGuigan & Dollins (1989) showed with electromyography that the tongue and lip muscles are activated in covert speech in the same way as during overt speech. An fMRI study by Wildgruber *et al.* (1996) showed primary motor cortex activation during covert speech. A recent study by Aziz-Zadeh *et al.* (2005) showed covert speech arrest after transient inactivation with repetitive transcranial magnetic stimulation (rTMS) over the left primary motor cortex and left BA44.

The above-mentioned presence in Broca's region of both hand and mouth motor representations is crucial not only for the evolution of language (Rizzolatti & Arbib 1998; Corballis 2002, 2004; Arbib 2005; Gentilucci & Corballis 2006), but also for its ontogeny. Developmental psychologists have shown the existence of a close relationship between the development of manual and oral motor skills. Goldin-Meadow (1999) proposed that speech production and speech-related hand gestures could be considered as outputs of the same process. Canonical babbling in children aged 6–8

months is accompanied by rhythmic hand movements (Masataka 2001). Hearing babies born to deaf parents display hand actions with a babbling-like rhythm. Manual gestures pre-date early development of speech in children, and predict later success even up to the two-word level (Iverson & Goldin-Meadow 2005).

It must be emphasized that the same intimate relationship between manual and oral language-related gestures persists in adulthood. Several pioneering works by Gentilucci and colleagues (Gentilucci 2003; Gentilucci *et al.* 2001, 2004a,b) have demonstrated a close relationship between speech production and the execution/observation of arm and hand gestures. In one of these studies (Gentilucci *et al.* 2004a), participants were required either to grasp and bring to the mouth fruits of different size like a cherry or an apple, or to observe the same actions performed by someone else, while simultaneously uttering the syllable 'ba'. The results showed that the second formant of the vowel 'a' (related to tongue position) increased when they executed or observed the act of bringing the apple (the larger object) to the mouth, or its pantomime, with respect to when they did the same with the cherry (the smaller object).

The execution/observation of the action of bringing an object to the mouth activates a mouth articulation posture probably related to food manipulation, which selectively influences speech production. This suggests that the system involved in speech production shares (and may derive from) the neural premotor circuit involved in the control of hand/arm actions.

In another related study (Gentilucci *et al.* 2004b), both adults and 6-year-old children were required to observe grasping and bringing to the mouth actions performed by others while uttering the syllable 'ba'. The results showed that the different observed actions influenced lip-shaping kinematics and voice formants. The observation of grasping influenced the first formant (which is related to mouth opening), while the observation of bringing to the mouth, as in the previous experiment, influenced the second formant of the voice spectrum, related to tongue position. It must be stressed that the effects on speech were greater in children. This study indicates that action observation induces the activation of the normally subsequent motor act in the observer; that is, mouth grasping when observing hand grasping and chewing when observing bringing to the mouth. This in turn affects speech production. As proposed by the authors of this study, this mechanism may have enabled the transfer from a primitive arm gesture communication system to speech. Given the stronger effects displayed by the children when compared with the adults, the same mechanism could be useful during speech learning in infancy.

In a very recent paper, Bernardis & Gentilucci (2006) asked participants to pronounce words (e.g. bye-bye, stop), to execute communicative arm gestures with the same meaning or to emit the two communication signals simultaneously. The results showed that the voice spectra of spoken words were reinforced by the simultaneous execution of the corresponding-in-meaning gesture when compared with those of word pronunciation alone. This was not observed when the gesture was meaningless. Conversely, pronouncing

words tended to inhibit the simultaneous execution of the gesture, as shown by the slowing down of the arm kinematics parameters. Comparable effects were not observed when pseudo-words were pronounced.

The results therefore showed that the word and the corresponding-in-meaning communicative gesture influenced each other when they were emitted simultaneously. The second formant in the voice spectra was higher when the word was pronounced together with the gesture. No modification in the second formant was observed when executing a meaningless arm movement, which nevertheless involved the same joints as the three meaningful gestures. Conversely, the second formant of a pseudo-word was not affected by the meaningful gestures.

Next, it was tested whether observing word pronunciation during gesture execution affected verbal responses in the same way as emitting the two signals. The voice spectra of words pronounced in response to simultaneously listening to and observing the speaker making the corresponding-in-meaning gesture were reinforced, just as they were by the simultaneous emission of the two communication signals.

The results of this elegant study seem to suggest that spoken words and symbolic communicative gestures are coded as a single signal by a unique communication system within the premotor cortex. The involvement of Broca's area in translating the representations of communicative arm gestures into mouth articulation gestures was recently confirmed by transient inactivation of BA44 with rTMS (Gentilucci *et al.* 2006). Since this brain region contains mirror neurons, it is most probable that through embodied simulation the communicative meaning of gestures is fused with the articulation of sounds required to express them in words.

7. EMBODIED SIMULATION AND LANGUAGE: MOTOR SIMULATION AT THE CONTENT LEVEL

The meaning of a sentence, regardless of its content, has been classically considered to be understood by relying on symbolic, amodal mental representations (Pylyshyn 1984; Fodor 1998). An alternative hypothesis, now more than 30 years old, assumes that the understanding of language relies on 'embodiment' (Lakoff & Johnson 1980, 1999; Lakoff 1987; Glenberg 1997; Barsalou 1999; Glenberg & Robertson 2000; Pulvermüller 1999, 2002, 2005; Gallese 2003c; Feldman & Naranayan 2004; Gallese & Lakoff 2005; Gentilucci & Corballis 2006).

According to the embodiment theory, for action-related sentences, the neural structures presiding over action execution should also play a role in understanding the semantic content of the same actions when verbally described. Empirical evidence shows this to be the case. Glenberg & Kaschak (2002) asked participants to judge if a read sentence was sensible or nonsense by moving their hand to a button, requiring movement away from the body (in one condition) or towards the body (in the other condition). Half of the sensible sentences described action towards the reader and half away. Readers responded faster to sentences describing actions whose direction was congruent with

the required response movement. This clearly shows that action contributes to sentence comprehension.

The most surprising result of this study, though, was that the same interaction between sentence movement direction and response direction was also found with abstract sentences describing transfer of information from one person to another, such as 'Liz told you the story' versus 'you told Liz the story'. These latter results extend the role of action simulation to the understanding of sentences describing abstract situations. Similar results were recently published by other authors (Borghi *et al.* 2004; Matlock 2004).

A prediction of the embodiment theory of language understanding is that when individuals listen to action-related sentences, their mirror neuron system should be modulated. The effect of this modulation should influence the excitability of the primary motor cortex, henceforth the production of the movements it controls. To test this hypothesis, we carried out two experiments (Buccino *et al.* 2005). In the first experiment, by means of single-pulse TMS, either the hand or the foot/leg motor areas in the left hemisphere were stimulated in distinct experimental sessions, while participants were listening to sentences expressing hand and foot actions. Listening to abstract content sentences served as a control. MEPs were recorded from hand and foot muscles. Results showed that MEPs recorded from hand muscles were specifically modulated by listening to hand action-related sentences, as were MEPs recorded from foot muscles by listening to foot action-related sentences.

In the second behavioural experiment, participants had to respond with the hand or the foot while listening to sentences expressing hand and foot actions when compared with abstract sentences. Coherently, with the results obtained with TMS, reaction times of the two effectors were specifically modulated by the effector-congruent heard sentences. These data show that processing sentences describing actions activates different sectors of the motor system, depending on the effector used in the listened action.

Several brain-imaging studies have shown that processing linguistic material in order to retrieve its meaning activates regions of the motor system congruent with the processed semantic content. Hauk *et al.* (2004) showed in an event-related fMRI study that silent reading of words referring to face, arm or leg actions led to the activation of different sectors of the premotor-motor areas that were congruent with the referential meaning of the read action words. Tettamanti *et al.* (2005) showed that listening to sentences expressing actions performed with the mouth, the hand and the foot produces activation of different sectors of the premotor cortex, depending on the effector used in the listened action-related sentence. These activated sectors correspond, albeit only coarsely, with those active during the observation of hand, mouth and foot actions (Buccino *et al.* 2001).

These data support the notion that the mirror neuron system is involved not only in understanding visually presented actions, but also in mapping acoustically presented action-related sentences. The precise functional relevance of the involvement of action embodied simulation for language understanding

remains unclear. One could speculate that such an involvement is purely parasitic, or, at best, reflects motor imagery induced by the upstream understanding process. The study of the spatio-temporal dynamic of language processing becomes crucial in settling this issue. Evoked readiness potential (ERP) experiments on silent reading of face-, arm- and leg-related words showed category-specific differential activations approximately 200 ms after word onset. Distributed source localization performed on stimulus-triggered ERPs showed different somatotopically arranged activation sources, with a strongest inferior frontal source for face-related words and a maximal superior central source for leg-related words (Pulvermüller *et al.* 2000).

This dissociation in brain activity patterns supports the idea of stimulus-triggered early lexico-semantic processes taking place within the premotor cortex. In order to control for a putative role of motor preparation processes in determining that effect, the same group of researchers carried out experiments in which the same response—a button press with the left index finger—was required for all words (Hauk & Pulvermüller 2004). The results showed a persistence of the early activation difference between face- and leg-related words, thus ruling out the motor preparation hypothesis. Pulvermüller *et al.* (2003) used magnetoencephalography to investigate the time course of cortical activation underlying the magnetic mismatch negativity elicited by hearing a spoken action-related word. The results showed that auditory areas of the left superior temporal lobe became active 136 ms after the information in the acoustic input was sufficient for identifying the word, and activation of the left inferior frontal cortex followed after an additional delay of 22 ms.

In sum, although these results are far from being conclusive on the effective relevance of the embodied simulation of action for language understanding, they show that simulation is specific, automatic and has a temporal dynamic compatible with such a function. More inactivation studies will be required to validate what at present is a little more than a plausible hypothesis.

8. EMBODIED SIMULATION, ACTION AND SYNTAX

I have reviewed in the previous sections empirical evidence demonstrating a consistent involvement of action and motor cortical circuits in various aspects of social cognition, including the processing of language. We should now frame what we have discussed so far about action, social cognition and language within an evolutionary perspective, and in doing so, introduce syntax.

Hauser *et al.* (2002) proposed to differentiate two domains within the language faculty: a 'narrow language faculty' (LFN), encompassing aspects that are specific to language, and a 'broad language faculty', supposedly inclusive of more general cognitive functions, not unique to humans, but shared with non-human animals. According to the same proposal, at the core of LFN is 'recursion', a specifically human computational mechanism at the basis of language grammar, which, nevertheless, might have evolved for functions other than language. The merit of this proposal in my opinion

lies in its greater evolutionary plausibility in comparison with alternative discontinuist views, like those positing a linguistic 'big-bang' out of which full-blown human language supposedly emerged (Bickerton 1995). It should be emphasized that even critics of the 'recursion-only hypothesis' applauded the merit of abandoning a monolithic view of language (see Pinker & Jackendoff 2005).

If embodied simulation is crucial in social cognition, language being the most distinctively human component of social cognition, syntax appears to be a crucial domain in which the relevance of embodied simulation for human social cognition can be tested. Syntax is a basic ingredient of the LFN, as defined by Hauser *et al.* (2002). According to the modular approach to syntax, syntactic processing is typically operated by a serial parsing encapsulated system, in which the initial phase of processing has access only to information about syntax. According to Fodor (1983, p. 77), '...to show that [the syntactic] system is penetrable (hence informationally unencapsulated), you would have to show that its processes have access to information that is not specified at any of the levels of representation that the language input system computes'.

Recent behavioural studies, though, show that the syntactic system *is* penetrable. Syntactic ambiguities are evaluated using non-linguistic constraints like real-world properties of referential context. Empirical research shows that humans continuously define linguistically relevant referential domains by evaluating sentence information against the situation-specific affordances. These affordances are not encoded as part of the linguistic representation of a word or phrase. Listeners use predicate-based information, like action goals, to anticipate upcoming referents. For example, a recent study by Chambers *et al.* (2004) shows that syntactic decisions about ambiguous sentences are affected by the number of referential candidates that can afford the action evoked by the verb in the unfolding sentence. These results suggest that even a key component of the supposed LFN is intimately intertwined with action and its embodied simulation.

A further evidence of the involvement of goal-related action with syntax comes from the fMRI studies, showing a clear relationship between the premotor system and the mapping of sequential events. Schubotz & von Cramon (2004) contrasted the observation of biological hand actions with that of abstract motion (movements of geometric shapes). In both conditions, 50% of the stimuli failed to attain the normally predictable end-state. The task of participants was to indicate whether the actions were performed in a goal-directed manner or not, and whether the abstract motions were performed regularly or not. Results showed that both conditions elicited significant activation within the ventral premotor cortex. In addition, the prediction of biological actions also activated BA44/45, which is part of the mirror neuron system. Schubotz & von Cramon (2004) concluded that their findings point to a basic premotor contribution to the representation or processing of sequentially structured events. This contribution appears to be even more specifically related to language, as the fMRI studies have shown selective

activation of premotor BA44 during the acquisition of artificial linguistic grammars characterized by long-distance, non-local syntactic dependencies (Tettamanti *et al.* 2002; Musso *et al.* 2003; see also Friederici 2004).

We said that the human language faculty is grounded in the unique ability to process hierarchically structured recursive sequences, configured as a phrase structure grammar (PSG). The human species is capable of mastering PSG, while other non-human primate species are confined to the use of much simpler finite state grammars (FSGs; see Hauser *et al.* 2002; Hauser & Fitch 2004). A recent fMRI study by Friederici *et al.* (2006) shows that the premotor sector of the inferior frontal gyrus, part of the mirror neuron system, is specifically activated during the processing of an artificial grammar bearing the PSG structure.

On the basis of all these results, it can be hypothesized that PSG is the computational output of a cortical premotor network originally evolved to control/represent the hierarchical structure of goal-related action. When in evolution, selective pressure led to the emergence of language, the same neural circuits doing computations to control the hierarchy of goal-related actions were 'exploited' to serve the newly acquired function of language syntax. A similar functional overlap between action and language acquisition is indeed evident during children's development, i.e. children parallel their capacity to master hierarchical complexity both in the domain of language and goal-related action (Greenfield 1991). My hypothesis can be easily tested with brain-imaging experiments. The prediction is that the opercular region of the inferior frontal gyrus should be activated by tasks involving the processing of complex, PSG-like hierarchical structures, both in the domain of action and language.

9. COGNITIVE CONTINUITY IN PRIMATES' SOCIAL COGNITION: THE NEURAL EXPLOITATION HYPOTHESIS

We are now in the position to better specify the wider implications of embodied simulation for social cognition, by formulating the neural exploitation hypothesis. The main claim is that key aspects of human social cognition are underpinned by neural exploitation; that is, the adaptation of sensory-motor-integrating brain mechanisms to serve new roles in thought and language, while retaining their original functions as well (see Gallese 2003c; Gallese & Lakoff 2005).

The execution of any complex coordinated action must make use of *at least* two brain sectors—the premotor and motor cortices, which are linked by reciprocal neural connections. The motor cortex controls individual synergies—relatively simple movements like extending and flexing the fingers, turning the wrist, flexing and extending the elbow, etc. The role of the premotor cortex is—not surprisingly—motor control, i.e. structuring such simple behaviours into coordinated motor acts, with the simple synergies performed at the right time, moving in the right direction, with the right force, for the right duration. This implies that the premotor cortex must provide a phase structure to actions and specify the right parameter values in the right phases. This information

must be conveyed from the premotor to the motor cortex by neural connections activating specific regions of the motor cortex. In addition, as epitomized by the mirror neuron system, the same premotor circuitry that governs motor control for action execution must govern the embodied simulation of the observed actions of others.

There is therefore a 'structuring' computational circuit within the premotor system that can function in two modes of operation. In the first mode, the circuit can structure action execution and/or action perception and imagination, with neural connections to motor effectors and/or other sensory cortical areas. In the second mode of operation, the same system is decoupled from its action execution/perception functions and can offer its structuring computations to non-sensory-motor parts of the brain (see Lakoff & Johnson 1999; Gallese & Lakoff 2005). As a result, the computational structure of the premotor system is applied, on the one hand, to master the hierarchical structure of language and, on the other hand, to 'abstract' domains, yielding 'abstract inferences'. According to this hypothesis, the same circuitry that controls how to move our body and enables our understanding of the action of others can, in principle, also structure language and abstract thought.

How can we reconcile the undisputable discontinuity among primate species in the capacity of processing complex recursive structures with the idea of cognitive continuity in primates' evolution of social cognition? My suggestion is that one important difference between humans and non-human primates could be the higher level of recursivity attained in our species—among many other neural systems—by the premotor cortex, of which the mirror neuron system is part. In fact, considering the impressive amount of evidence reviewed above, the premotor system is probably one of the most important brain regions where this evolutionary process might have taken place. The hypothesis I put forward is that the quantitative difference in computational power and degree of recursivity attained by the human brain—and, in particular, by the mirror neuron system—with respect to the brains of non-human primates could produce a qualitative leap forward in social cognition.

However, the computational divide between humans and other primates is probably not the only explanation. A second consideration must be added. The evolution of social cognition should not be conceived like a monotonic function, with a strict correlation between the chronological position a species occupies in phylogeny and its level of social 'cognitive smartness'. Hare & Tomasello (2005) show that dogs exhibit social communicative skills in tasks where apes fail, like finding food on the basis of human communicative gestures like pointing or gaze cues. These authors suggest that the remarkable social communicative skills displayed by dogs could be the outcome of their domestication process. This would represent a case of convergent evolution with humans, in which the initial selection of strictly speaking 'non-cognitive', emotional traits like tameness could have played a crucial bootstrapping role. If Hare & Tomasello (2005) are right, then one could argue that

the specific social cognitive endowments of our species are the evolutionary outcome of the selection of mechanisms that are not intrinsically cognitive or, at the very least, certainly not mind-reading specific.

The appeal of the present hypothesis consists in its parsimony. Embodied simulation and its neural underpinnings may well fall short of providing a thorough account of what is implied in our sophisticated social cognitive skills. However, I believe that the evidence presented here indicates that embodied mechanisms involving the activation of the premotor system, of which the mirror neuron system is a part, do play a major role in social cognition.

10. CONCLUSIONS

Our sophisticated mind-reading abilities probably involve the activation of large regions of our brain, certainly larger than a putative and domain-specific theory of mind module. My point is that these brain sectors do encompass the premotor system and, in particular, the mirror neuron system. The social use of language is one of the most powerful cognitive tools to understand others' minds. Embodied simulation mechanisms are involved in language processing, and might also be crucial in the course of the long learning process children require to become fully competent in how to use folk psychology. This learning process greatly benefits from the repetitive exposure to the narration of stories about the actions of various characters (for a putative role of narrative practices in the development of a competent use of folk psychology, see Hutto 2004).

As suggested by Arciero (2006), to imbue words with meaning requires a fusion between the articulated sound of words and the shared meaning of action. Embodied simulation does exactly that. Furthermore, and most importantly, embodied simulation and the mirror neuron system underpinning it provide the means to share communicative intentions, meaning and reference, thus granting the parity requirements of social communication (Tomasello *et al.* 2005).

As I have argued elsewhere (Gallese 2006; Gallese & Umiltà 2006), the automatic translation of the folk-psychology-inspired 'flow charts' into encapsulated brain modules, specifically adapted to mind-reading abilities, should be carefully scrutinized. Language can typically play ontological tricks by means of its 'constitutiveness'; that is, its capacity to give an apparent ontological status to the concepts words embody (Bruner 1986, p. 64). Space can provide an illuminating example of how our language-based definitions do not necessarily translate into real entities in the brain. Space, although unitary when examined introspectively, is not represented in the brain as a single multipurpose map. There is no central processing unit for space in our brain to support the unitary idea of it that humans entertain. On the contrary, in the brain there are numerous spatial maps (see Rizzolatti *et al.* 1997). The same might be true for our language-mediated definition of what it means to mind-read, namely the employment of the cognitive tools of folk psychology. We can do better than merely looking for the brain location of intentions, beliefs and desires as

such. A more promising and potentially fruitful strategy lies in the comparative study of the role played in social cognition by the premotor system of primate brains.

This work was supported by MIUR (Ministero Italiano dell'Istruzione, dell'Università e della Ricerca), and also this work, being a part of the European Science Foundation EUROCORES Programme OMLL, was supported by the funds to V.G. from the Italian C.N.R.

REFERENCES

- Arbib, M. A. 2005 From monkey-like action recognition to human language: an evolutionary framework for neuro-linguistics. *Behav. Brain Sci.* **28**, 105–168. (doi:10.1017/S0140525X05000038)
- Arciero, G. 2006 *Sulle tracce disé*. Milano, Italy: Bollati-Boringhieri.
- Aziz-Zadeh, L., Cattaneo, L., Rochat, M. & Rizzolatti, G. 2005 Covert speech arrest induced by rTMS over both motor and nonmotor left hemisphere frontal sites. *J. Cogn. Neurosci.* **17**, 928–938. (doi:10.1162/0898929054021157)
- Barsalou, L. W. 1999 Perceptual symbol systems. *Behav. Brain Sci.* **22**, 577–609. (doi:10.1017/S0140525X99002149)
- Bernardis, P. & Gentilucci, M. 2006 Speech and gesture share the same communication system. *Neuropsychologia* **44**, 178–190. (doi:10.1016/j.neuropsychologia.2005.05.007)
- Bickerton, D. 1995 *Language and human behavior*. Seattle, WA: University of Washington Press.
- Bookheimer, S. 2002 Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* **25**, 151–188. (doi:10.1146/annurev.neuro.25.112701.142946)
- Borghi, A. M., Glenberg, A. M. & Kaschak, M. P. 2004 Putting words in perspective. *Mem. Cognit.* **32**, 863–873.
- Bruner, J. 1986 *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Bruner, J. 1990 *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Buccino, G. *et al.* 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404. (doi:10.1046/j.1460-9568.2001.01385.x)
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A. & Rizzolatti, G. 2004a Neural circuits involved in the recognition of actions performed by nonconspicuous: an fMRI study. *J. Cogn. Neurosci.* **16**, 114–126. (doi:10.1162/089892904322755601)
- Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H.-J. & Rizzolatti, G. 2004b Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* **42**, 323–334. (doi:10.1016/S0896-6273(04)00181-3)
- Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V. & Rizzolatti, G. 2005 Listening to action-related sentences modulates the activity of the motor system: a combined TMS and behavioral study. *Cogn. Brain Res.* **24**, 355–363. (doi:10.1016/j.cogbrainres.2005.02.020)
- Buxbaum, L. J., Kyle, K. M. & Menon, R. 2005 On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cogn. Brain Res.* **25**, 226–239. (doi:10.1016/j.cogbrainres.2005.05.014)
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E. & Haggard, P. 2005 Action observation and acquired motor skills: an fMRI study with expert dancers. *Cerebral Cortex* **15**, 1243–1249. (doi:10.1093/cercor/bhi007)
- Chambers, C. G., Tanenhaus, M. K. & Magnuson, J. S. 2004 Actions and affordances in syntactic ambiguity resolution. *J. Mem. Lang.* **30**, 687–696.

- Corballis, M. C. 2002 *From hand to mouth: the origins of language*. Princeton, NJ: Princeton University Press.
- Corballis, M. C. 2004 FOXP2 and the mirror system. *Trends Cogn. Sci.* **8**, 95–96. (doi:10.1016/j.tics.2004.01.007)
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. 1992 Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180. (doi:10.1007/BF00230027)
- Fadiga, L., Craighero, L., Buccino, G. & Rizzolatti, G. 2002 Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* **15**, 399–402. (doi:10.1046/j.0953-816x.2001.01874.x)
- Feldman, J. & Narayanan, S. 2004 Embodied meaning in a neural theory of language. *Brain Lang.* **89**, 385–392. (doi:10.1016/S0093-934X(03)00355-9)
- Ferrari, P. F., Gallese, V., Rizzolatti, G. & Fogassi, L. 2003 Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* **17**, 1703–1714. (doi:10.1046/j.1460-9568.2003.02601.x)
- Ferrari, P. F., Maiolini, C., Addessi, E., Fogassi, L. & Visalberghi, E. 2005 The observation and hearing of eating actions activates motor programs related to eating in macaque monkeys. *Behav. Brain Res.* **161**, 95–101. (doi:10.1016/j.bbr.2005.01.009)
- Fodor, J. 1983 *The modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, J. 1998 *Concepts*. Oxford, UK: Oxford University Press.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **302**, 662–667. (doi:10.1126/science.1106138)
- Friederici, A. D. 2004 Processing local transitions versus long-distance syntactic hierarchies. *Trends Cogn. Sci.* **8**, 245–247. (doi:10.1016/j.tics.2004.04.013)
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I. & Anwander, A. 2006 The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl Acad. Sci. USA* **103**, 2458–2463. (doi:10.1073/pnas.0509389103)
- Gallese, V. 2001 The “shared manifold” hypothesis: from mirror neurons to empathy. *J. Conscious. Stud.* **8**, 33–50.
- Gallese, V. 2003a The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. B* **358**, 517–528. (doi:10.1098/rstb.2002.1234)
- Gallese, V. 2003b The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* **36**, 171–180. (doi:10.1159/000072786)
- Gallese, V. 2003c A neuroscientific grasp of concepts: from control to representation. *Phil. Trans. R. Soc. B* **358**, 1231–1240. (doi:10.1098/rstb.2003.1315)
- Gallese, V. 2005a Embodied simulation: from neurons to phenomenal experience. *Phenomenol. Cogn. Sci.* **4**, 23–48. (doi:10.1007/s11097-005-4737-z)
- Gallese, V. 2005 “Being like me”: self-other identity, mirror neurons and empathy. In *Perspectives on imitation: from cognitive neuroscience to social science*, vol. 1 (eds S. Hurley & N. Chater), pp.101–118. Cambridge, MA: MIT Press.
- Gallese, V. 2006 Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Cogn. Brain Res.* **1079**, 15–24.
- Gallese, V. & Goldman, A. 1998 Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* **12**, 493–501. (doi:10.1016/S1364-6613(98)01262-5)
- Gallese, V. & Lakoff, G. 2005 The brain’s concepts: the role of the sensory-motor system in reason and language. *Cogn. Neuropsychol.* **22**, 455–479.
- Gallese, V. & Umiltà, M. A. 2006 Cognitive continuity in primate social cognition. *Biol. Theory* **1**, 25–30. (doi:10.1162/biot.2006.1.1.25)
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)
- Gallese, V., Fogassi, L., Fadiga, L. & Rizzolatti, G. 2002 Action representation and the inferior parietal lobule. In *Attention and performance XIX* (eds W. Prinz & B. Hommel), pp. 247–266. Oxford, UK: Oxford University Press.
- Gallese, V., Keysers, C. & Rizzolatti, G. 2004 A unifying view of the basis of social cognition. *Trends Cogn. Sci.* **8**, 396–403. (doi:10.1016/j.tics.2004.07.002)
- Gentilucci, M. 2003 Grasp observation influences speech production. *Eur. J. Neurosci.* **17**, 179–184. (doi:10.1046/j.1460-9568.2003.02438.x)
- Gentilucci, M. & Corballis, M. C. 2006 From manual gesture to speech: a gradual transition. *Neurosci. Biobehav. Rev.* **30**, 949–960. (doi:10.1016/j.neubiorev.2006.02.004)
- Gentilucci, M., Benuzzi, F., Gangitano, M. & Grimaldi, S. 2001 Grasp with hand and mouth: a kinematic study on healthy subjects. *J. Neurophysiol.* **86**, 1685–1699.
- Gentilucci, M., Santunione, P., Roy, A. C. & Stefanini, S. 2004a Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *Eur. J. Neurosci.* **19**, 190–202. (doi:10.1111/j.1460-9568.2004.03104.x)
- Gentilucci, M., Stefanini, S., Roy, A. C. & Santunione, P. 2004b Action observation and speech production: study on children and adults. *Neuropsychologia* **42**, 1554–1567. (doi:10.1016/j.neuropsychologia.2004.03.002)
- Gentilucci, M., Bernardis, P., Crisi, G. & Volta, R. D. 2006 Repetitive transcranial magnetic stimulation of Broca’s area affects verbal responses to gesture observation. *J. Cogn. Neurosci.* **18**, 1059–1074. (doi:10.1162/jocn.2006.18.7.1059)
- Glenberg, A. M. 1997 What memory is for. *Behav. Brain Sci.* **20**, 1–19. (doi:10.1017/S0140525X97000010)
- Glenberg, A. M. & Kaschak, M. P. 2002 Grounding language in action. *Psychon. Bull. Rev.* **9**, 558–565.
- Glenberg, A. M. & Robertson, D. A. 2000 Symbol grounding and meaning: a comparison of high-dimensional and embodied theories of meaning. *J. Mem. Lang.* **43**, 379–401. (doi:10.1006/jmla.2000.2714)
- Goldin-Meadow, S. 1999 The role of gesture in communication and thinking. *Trends Cogn. Sci.* **3**, 419–429. (doi:10.1016/S1364-6613(99)01397-2)
- Greenfield, P. M. 1991 Language, tools, and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**, 531–550.
- Hare, B. & Tomasello, M. 2005 Human-like social skills in dogs? *Trends Cogn. Sci.* **9**, 439–444. (doi:10.1016/j.tics.2005.08.010)
- Hauk, O. & Pulvermüller, F. 2004 Neurophysiological distinction of action words in the fronto-central cortex. *Hum. Brain Mapp.* **21**, 191–201. (doi:10.1002/hbm.10157)
- Hauk, O., Johnsrude, I. & Pulvermüller, F. 2004 Somatotopic representation of action words in human motor and premotor cortex. *Neuron* **41**, 301–307. (doi:10.1016/S0896-6273(03)00838-9)
- Hauser, M. D. & Fitch, W. T. 2004 Computational constraints on syntactic processing in a non-human primate. *Science* **303**, 377–380. (doi:10.1126/science.1089401)
- Hauser, M. D., Chomsky, N. & Fitch, W. T. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579. (doi:10.1126/science.298.5598.1569)
- Hutto, D. H. 2004 The limits of spectatorial folk psychology. *Mind Lang.* **19**, 548–573.

- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. & Rizzolatti, G. 1999 Cortical mechanisms of human imitation. *Science* **286**, 2526–2528. (doi:10.1126/science.286.5449.2526)
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. & Rizzolatti, G. 2005 Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* **3**, 529–535. (doi:10.1371/journal.pbio.0030079)
- Iverson, J. M. & Goldin-Meadow, S. 2005 Gesture paves the way for language development. *Psychol. Sci.* **16**, 367–371. (doi:10.1111/j.0956-7976.2005.01542.x)
- Jacob, P. & Jeannerod, M. 2005 The motor theory of social cognition: a critique. *Trends Cogn. Neurosci.* **9**, 21–25. (doi:10.1016/j.tics.2004.11.003)
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V. & Rizzolatti, G. 2002 Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848. (doi:10.1126/science.1070311)
- Lakoff, G. 1987 *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago, IL; London, UK: University of Chicago Press.
- Lakoff, G. & Johnson, M. 1980 *Metaphors we live by*. Chicago, IL; London, UK: University of Chicago Press.
- Lakoff, G. & Johnson, M. 1999 *Philosophy in the flesh*. New York, NY: Basic Books.
- Masataka, N. 2001 Why early linguistic milestones are delayed in children with Williams syndrome: late onset of hand banging as a possible rate-limiting constraint on the emergence of canonical babbling. *Devel. Sci.* **4**, 158–164. (doi:10.1111/1467-7687.00161)
- Matlock, T. 2004 Fictive motion as cognitive simulation. *Mem. Cognit.* **32**, 1389–1400.
- McGuigan, F. J. & Dollins, A. B. 1989 Patterns of covert speech behavior and phonetic coding. *Pavlovian J. Biol. Sci.* **24**, 19–26.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C. & Weillier, C. 2003 Broca's area and the language instinct. *Nat. Neurosci.* **6**, 774–781. (doi:10.1038/nn1077)
- Nishitani, N., Schürmann, M., Amunts, K. & Hari, R. 2005 Broca's region: from action to language. *Physiology* **20**, 60–69. (doi:10.1152/physiol.00043.2004)
- Onishi, K. H. & Baillargeon, R. 2005 Do 15 months-old understand false beliefs? *Science* **308**, 255–258. (doi:10.1126/science.1107621)
- Paukner, A., Anderson, J. R., Borelli, E., Visalberghi, E. & Ferrari, P. F. 2005 Macaques (*Macaca nemestrina*) recognize when they are being imitated. *Biol. Lett.* **1**, 219–222. (doi:10.1098/rsbl.2004.0291)
- Pinker, S. & Jackendoff, R. 2005 The faculty of language: what's special about it? *Cognition* **95**, 201–236. (doi:10.1016/j.cognition.2004.08.004)
- Pulvermüller, F. 1999 Word in the brain's language. *Behav. Brain Sci.* **22**, 253–336. (doi:10.1017/S0140525X9900182X)
- Pulvermüller, F. 2002 *The neuroscience of language*. Cambridge, UK: Cambridge University Press.
- Pulvermüller, F. 2005 Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* **6**, 576–582. (doi:10.1038/nrn1706)
- Pulvermüller, F., Härle, M. & Hummel, F. 2000 Neurophysiological distinction of verb categories. *Neuroreport* **11**, 2789–2793.
- Pulvermüller, F., Shtyrov, Y. & Ilmoniemi, R. J. 2003 Spatio-temporal patterns of neural language processing: an MEG study using minimum-norm current estimates. *Neuroimage* **20**, 1020–1025. (doi:10.1016/S1053-8119(03)00356-2)
- Pylyshyn, Z. W. 1984 *Computation and cognition: toward a foundation for cognitive science*. Cambridge, MA: MIT Press.
- Rizzolatti, G. & Arbib, M. A. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194. (doi:10.1016/S0166-2236(98)01260-0)
- Rizzolatti, G. & Craighero, L. 2004 The mirror neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)
- Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. 1996 Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141. (doi:10.1016/0926-6410(95)00038-0)
- Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. 1997 The space around us. *Science* **277**, 190–191. (doi:10.1126/science.277.5323.190)
- Rizzolatti, G., Fogassi, L. & Gallese, V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Neurosci. Rev.* **2**, 661–670. (doi:10.1038/35090060)
- Schubotz, R. I. & von Cramon, D. Y. 2004 Sequences of abstract non-biological stimuli share ventral premotor cortex with action observation and imagery. *J. Neurosci.* **24**, 5467–5474. (doi:10.1523/JNEUROSCI.1169-04.2004)
- Subiaul, F., Cantlon, J. F., Holloway, R. L. & Terrace, H. S. 2004 Cognitive imitation in rhesus macaque. *Science* **305**, 407–410. (doi:10.1126/science.1099136)
- Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S. & Weniger, D. 2002 Neural correlates for the acquisition of natural language syntax. *Neuroimage* **17**, 700–709. (doi:10.1016/S1053-8119(02)91201-2)
- Tettamanti, M. et al. 2005 Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* **17**, 273–281. (doi:10.1162/0898929053124965)
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. 2005 Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691. (doi:10.1017/S0140525X05000129)
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. & Rizzolatti, G. 2001 "I know what you are doing": a neurophysiological study. *Neuron* **32**, 91–101.
- Watkins, K. E. & Paus, T. 2004 Modulation of motor excitability during speech perception: the role of Broca's area. *J. Cogn. Neurosci.* **16**, 978–987. (doi:10.1162/0898929041502616)
- Watkins, K. E., Strafella, A. P. & Paus, T. 2003 Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* **41**, 989–994. (doi:10.1016/S0028-3932(02)00316-0)
- Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B. & Grodd, W. 1996 Functional lateralization of speech production at primary motor cortex: a fMRI study. *NeuroReport* **7**, 2791–2795. (doi:10.1097/00001756-199611040-00077)
- Wilson, S. M., Saygin, A. P., Sereno, M. I. & Iacoboni, M. 2004 Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* **7**, 701–702. (doi:10.1038/nn1263)